

Critiquing: A methodology to extract C2 expertise

Track: C2 Decision Making and Cognitive Analysis

Dr. Janet E. Miller

Air Force Research Laboratory/HECA,

Wright-Patterson AFB, OH 45433

Voice: 937-255-8805; Fax: 937-255-9198

Janet.Miller3@wpafb.af.mil

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE JUN 2003		2. REPORT TYPE		3. DATES COVERED 00-00-2003 to 00-00-2003	
4. TITLE AND SUBTITLE Critiquing: A methodology to extract C2 expertise				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Research Laboratory (HECA), Wright Patterson AFB, OH, 45433				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 34	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Critiquing: A methodology to extract C2 expertise

Dr. Janet E. Miller
AFRL/HECA
Wright-Patterson AFB, OH 45433
Janet.Miller3@wpafb.af.mil

ABSTRACT

The higher a military person goes in the ranks, the sooner the transfer, meaning the commanders who make the highest level decisions have been in a particular assignment the shortest amount of time. To acquire and maintain the maximum level of efficiency and productivity as quickly as possible, the expertise of commanders must be understood so that appropriate decision support can be developed. Cognitive task analysis methods, including knowledge elicitation, can be applied to uncover the information about how a domain practitioner works. Such information on practitioner behavior provides valuable insight and data for incorporation into models used for training and development of automated support. However, knowledge elicitation methods face a number of challenges such as grounding in context, limited accessibility to experts and tasks, being labor intensive and time consuming, and difficulties with repeatability. In this study, a critiquing methodology was investigated in its ability to address these challenges. This baseline study involved a novice intelligence analyst performing a basic analysis task. Then, six experts with various backgrounds critiqued the novice's process. The results suggest that the critiquing method addresses the challenges of knowledge elicitation methods and can be applied to understanding command and control.

INTRODUCTION

One given fact in the military is that the soldier, the sailor, and the airman will be transferred. The higher a military person goes in the ranks, the sooner the transfer, which means the commanders who make the highest level decisions have been in a particular organization and assignment the shortest amount of time. This method of continually transferring military members to new assignments on the average of every three years positively contributes to rounding out the knowledge and experience of the military person. However, each particular assignment, especially command and control, requires specific expertise beyond general knowledge. Even less frequently encountered are true direct military conflict assignments that have an inordinate level of dynamic, stressful, and risky challenges. To acquire and maintain the maximum level of efficiency and productivity in a particular assignment, including high-risk assignments such as Joint Forces Air Component Commander (JFACC), as quickly as possible, the

expertise of commanders must be understood so that appropriate decision support, including training, can be developed. Cognitive task analysis (CTA) reveals those specific facets of expertise that domain practitioners, such as commanders, utilize including the information requirements for decision making. Therefore, cognitive task analysis needs to be used with commanders to understand their decision challenges and problems so that appropriate decision support can be developed to aid in current decision making and to fortify against lost decision rationale and understanding, especially after personnel turnover.

Cognitive task analysis is part of the field of cognitive engineering that looks at knowledge, strategies, and goal structures that underlie observable work performance. CTA methodologies are used to discover the basis of expertise that domain practitioners use to perform tasks, but are unable to reliably articulate when asked directly. Specifically, they identify ineffective strategies that lead to poor performance (i.e., a model of mistakes that “novices” make), as well as adaptive strategies that have been developed by highly skilled practitioners to cope with task demands (i.e., a model of “expert” performance). A variety of these approaches have been developed for gaining insight to the performance of highly skilled practitioners. They range from field studies where practitioners are observed in actual work settings (Hutchins, 1995) to critical incident analyses (Klein, Calderwood, and Clinton-Cirocco, 1986) to analyses of performance under simulated conditions (Sarter and Woods, 2000) to observation of practitioner performance under highly controlled conditions (Lesgold *et al.*, 1988). Each of these methods has advantages and disadvantages, and can be combined to obtain converging evidence in order to “bootstrap” an understanding of the sources of expertise in a domain (Potter *et al.*, 2000).

A traditional beginning to cognitive task analysis is to gain some understanding about a domain through research, reading documentation, and other indirect means. Then direct knowledge acquisition methods are used to uncover information about how a practitioner works in a field of practice, including the work done 'in the head,' from the actual practitioner. These direct methods, known as knowledge elicitation methods, include a variety of interview techniques and different ways to observe task performance. Although undoubtedly any of the knowledge elicitation methods would reveal some knowledge about a domain, there are considerations on choosing the technique most likely to be effective for a particular domain. Considerations include whether the technique can authentically capture information about the domain, whether the results of the technique are commensurate with the expended resources, whether the resulting products are sufficiently documented, and whether the method generates ideas for domain support (Potter *et al.*, 2000). Good techniques encourage shared communication between the investigator and the practitioner, facilitate an unconstrained expression of knowledge, provide good data quality, and are able to be replicated (McNeese *et al.*, 1995; Hoffman, Crandall and Shadbolt, 1998).

Eliciting knowledge from experts is not easy. For one, careful thought must be given to planning and developing the details of the method to be applied to ensure the elicited knowledge is grounded in context. Another issue is potential difficulties gaining access

to both experts and to meaningful tasks. Access to experts can be an issue if there are only a few experts performing the task of interest, if there are physical access challenges, such as being in a restricted area, or if the experts are reluctant to be evaluated because of possible repercussions. Access to meaningful tasks can be an issue if the meaningful events occur at a low frequency, there is unpredictability of target events, or there are physical access limitations. For a third, the process is laborious and time consuming. Hours are required not only for developing the details of the session for eliciting knowledge but also to process the extracted information into a useable form. A last issue is repeatability, which supports making sense of the gathered information by allowing comparisons between sessions.

Command and control (C2) presents a unique domain for eliciting knowledge. For one, there are only so many C2 positions and while there are some similarities in the tasks performed, their differences are greater than their similarities. This situation contributes to both interesting task and expert access issues. Then, dynamic, critical C2 positions such as the JFACC only exist during conflict times when the commander is extremely occupied with the job at hand. Opportunity to capture expertise on particular cognitively demanding C2 tasks under these conditions is therefore extremely limited. Finally, repeatability and grounding in context are challenges as the context of conflicts has many variables that must be considered by the commander, including cultural and behavioral aspects of the country in focus.

Interviews and performance observations are knowledge elicitation methods that can address the challenges in different ways and to different degrees.

- Grounding in Context:

Interviews provide an opportunity for a domain practitioner to retrospectively tell how he performs work. This means that the practitioner is not physically grounded in context and that a definite challenge exists to have the practitioner be mentally grounded. The CTA investigator needs to have deep understanding of the work, which requires extensive pre-session work, so that the practitioner is provided realistic, stimulating questions to reveal his knowledge. The interviews focus on telling stories that can bring out interesting knowledge that the practitioner might not even realize is interesting. As the questions guide the session, the revealed knowledge is based on what the investigator has predetermined will be interesting. Performance observations are prospective as the practitioner is revealing what he is about to do as he shows what he does. In observations, then, the practitioner is both physically and mentally grounded in context as he is doing realistic work. Being bound in a realistic context allows for natural triggering of domain knowledge and promoting accurate verbalizations. Observations require that the CTA investigator has prior knowledge of the work to be prepared to capture interesting events. But as the session is not guided, the investigator does not need extensive prior knowledge on what will be interesting, which can reduce the pre-session planning. Another advantage of having a less guided session is that unexpected opportunities may naturally arise that the investigator can capture.

- Access Issues:

Physical access issues can be addressed with interviews, as the CTA investigator does not necessarily have to be in the practitioner's work area for the interview session.

Performance observation using simulations can also overcome physical access issues. Meaningful task access issues are addressed with well-planned interviews by directly asking about such events. Performance observations do not ensure access to meaningful events. If an unexpected, interesting event does occur, though, the investigator is available to gather data. If the performance observation is a simulation, the simulation can be based on intentional interesting events, such as a crisis.

- Laborious and Time Consuming:

With the interview technique, the questions provide a framework for the CTA investigator to know ahead of time when interesting information will be provided which can decrease the time to process the extracted information. While performance observations might require less time in the planning phase, the time to process the extracted information is increased, as the investigator has not predetermined what will be interesting.

- Repeatability:

Interviews allow repeatability. Observation of simulations supports repeatability but repeatability of actual work tasks is dependent upon the domain.

In summary, interview elicitations and performance observations have different strengths and weaknesses. The strengths of interviews are they overcome some of the expert and task access issues, they are less laborious and time consuming in processing the gathered data, and they allow repeatability. The weaknesses of interviews are they are not fully grounded in context, the practitioner is retrospectively telling how he does a task, they require more pre-session understanding to ensure interesting and relevant issues are discussed, and they do not provide for flexibility in gathering data on unexpected events. The strengths of performance observations are they are grounded in physical and mental context, the practitioner is prospectively showing and discussing how he does a task, they require less pre-session planning, and they allow the investigator to take advantage of unexpected events. The weaknesses of performance observations are they are laborious and time consuming to process the gathered data, they may have physical and task access issues, and they do not necessarily allow repeatability.

Interviews and observation methods have been useful in revealing how practitioners work in various domains and can be used in C2. However, developing another methodology that combines the strengths and addresses the weaknesses of the two methods would be beneficial to the field of CTA and can help address the particular challenges of capturing C2 expertise. This paper discusses a methodology that proposes to address that goal, critiquing by an expert. 'Expert' is used as shorthand for an experienced, knowledgeable practitioner, which is a more accurate categorization of the practitioners who participated in this study. A true expert is a practitioner who is distinguished or brilliant in the field, highly regarded by peers, whose judgments are uncommonly accurate and reliable, whose performance shows consummate skill and economy of effort and who can deal effectively with rare or 'tough' cases (Hoffman, *et al.*, 1995). 'Novice' is similarly used as shorthand for a person who is new to a domain although he may have skills that translate from prior experience in other domains.

Figure 1 illustrates some of the ways that interviewing, observation, and critiquing elicitation methods differ. With interviewing, the expert tells information to the investigator, often using stories based on past cases and experiences. While being questioned by the investigator, the expert as storyteller might reveal not only how he handled a particular case, but also have further comments on how that case changed his later work practices. The value of the results from interviews is mainly a function of the probing skill of the investigator and how well the interviewee understands what type of data is sought. With direct observation, the investigator watches the expert as a domain practitioner performing actual or simulated tasks. The value of the results is mainly a function of how realistic the scenarios and performances are and how well the observed events stress the cognitive system in order to reveal leverage points for improvement. With critiquing, the expert evaluates the performance of another practitioner. This sets up a situation where the domain expert is not concerned about his performance evaluation. In addition, the technique relies less on the investigator's probing skills and domain knowledge. The value of the results is mainly a function of the ability of the investigator to analyze the diverse data that is collected, which is of varying reliability and validity and not always structured by a story or scenario.

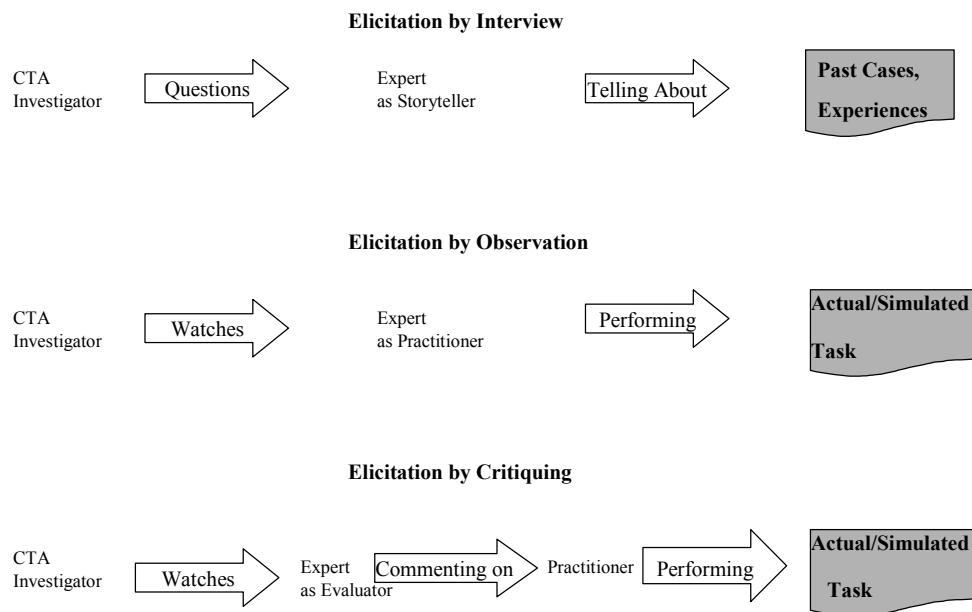


Figure 1. Comparison of Knowledge Elicitation Methods

Critiquing of a practitioner by an expert somewhat relates to the research on expert critiquing systems but is at the same time very different. Expert critiquing systems

(Fischer, *et al*, 1991), or critics, are a class of expert systems that receive as input the statement of the problem and the user-proposed solution. Likewise, critiquing as a CTA methodology is done by having an expert practitioner be presented with the task of interest and the novice's strategy. Automated critics recognize possible errors, form persuasive criticisms against what might be a chosen sub-optimal proposed solution, and adapt to the situation and any lessons learned. Critics are silent if no error is identified. Critiquers for the CTA methodology are encouraged to verbalize on possible errors as well as good actions taken by the novice and to make comments about what they would have done had they been working the task.

METHODOLOGY

As a baseline study into critiquing, this study was done in the military intelligence analysis domain involving a newly assigned officer performing a basic domain task. Then, the task and a presentation of the newly assigned officer's performance on the task were presented to six experienced domain practitioners for critiquing. The domain practitioners' task was to critique the newly assigned officer's performance, a task that is similar to the normal mentoring done by seasoned officers. The newly assigned officer's performance was used as a cue to the experts for discussion on what the experts consider correct and incorrect actions and to comment on what they would have done had they been working the domain task based on past experiences. A second novice was used for method verification.

The interest for this study was two-fold: to explore critiquing as a knowledge elicitation methodology and to investigate the domain of intelligence analysis with a focus on revealing insights that might be used for developing support models. The breadth of intelligence analysis work includes analyzing various types of media, such as physical artifacts and photography as well as text, and reporting in various ways from short briefings to long reports. To address both the methodology and domain interests this study was limited to investigating the intelligence analysts' basic process on textual data of information selection, corroboration/resolution, and story construction. The specific task statement was: "In 1996, the European Space Agency lost a satellite during the first qualification launch of new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were." This task was chosen after consultation with domain experts because the task was very representative of actual work accomplished in military intelligence analysis. The available electronic data set contained about 2000 text documents and had enough information to provide a detailed answer to the written task. Approximately 60% were 'on target' in that they contained information that could directly add to answering the task. Approximately 35% gave context but did not contain information specific to the task. The last 5% were irrelevant to the task.

First, a novice was given a task to perform and asked to talk aloud during the task. The task was to answer specific questions by using information in the provided electronic data set. The novice's process of developing a response was captured by video and

audio taping with his consent. To present the novice's process to the experts, the comments made by the novice during the analysis were transcribed into a script. In addition, the notes which the novice made during the analysis were recreated into successive PowerPoint slides to show the development of the note artifact over time. Finally, screen shots were made of his query responses and hard copies were made of the documents that he reviewed. The slides, screenshots and documents were shown at coordinated times with the script. A second novice also performed the task and the same data was collected to verify results. Both novices were recently commissioned U.S. Air Force second lieutenants. The first had received a B.S. in physics and the second novice had a B.S. in aerospace engineering. Neither had done intelligence analysis and the only task that either had performed which was similar to the study task was research in the academic environment.

Six experienced analysts (referred to as E5 to E10 because there were four pilot study participants) participated in this study and had been nominated by their organization to participate as representative experts. As can be seen in Table 1, their backgrounds varied with respect to years as an analyst, prior knowledge and educational background. Prior knowledge refers to how informed the experts were of the particular satellite launch incident prior to the critiquing session. Two of the analysts had no knowledge about the launch accident or the particular task prior to their critiquing session. Two of the analysts did not have knowledge about the incident from their work experience but were given the task statement and the documents that the first novice used in answering the task statement prior to their critiquing session. The final two analysts had both knowledge of the incident from their work experience and were given the task statement and the novice's documents prior to their critiquing session.

When each expert reported for his session, the problem task was read to the expert. The experts were asked to think aloud as the novice's process was revealed. Then, the script of the first novice's process was read with pauses to encourage commenting. As the script was read, the PowerPoint viewgraphs of the note artifact, screen shots of query responses and hard copies of the retrieved documents were shown. The experts' comments about the novice's process were captured on audio and videotape. The two experts who had full prior knowledge also critiqued the presentation of the second novice's process.

The data to be analyzed for the study was the experts' protocols from critiquing the novices. This data was iteratively analyzed using a process tracing methodology (Woods, 1993). First, a set of detailed protocols from each expert participant's process for the primary novice was constructed. The protocols were broken, by participant, into the frameworks of information selection, corroboration/resolution, and story construct. A comparison of the results of the primary novice was made to the results from the second novice to substantiate the findings.

Table 1. Background Information on Experienced Analysts

Participant	Gender	Years as Analyst	Prior Knowledge	University Degrees
E5	M	8	None	BSEE MSEE MS in Studies of Futures
E6	M	21	None	BS in Aero
E7	M	12.5	Full	BS in Aero, MA in Strategic Intelligence
E8	M	18	Partial	BA in Business Mgt
E9	M	32	Partial	BS in Physics BS in Math MS in Nuclear Physics
E10	M	12	Full	BA in Psychology MA in Intern'l Relations

RESULTS

Analysis of gathered information revealed that the three main tasks of information selection, corroboration/resolution and story construction could be broken down further into other subtasks. Information selection includes problem definition, applying a search strategy, and evaluating query results. Corroboration and resolution includes building a hypothesis or hypotheses against which to judge information and evaluating information. Story construction involves deciding on the best hypothesis and putting the story together. This refined breakdown was used for analysis of trends.

Tables 2-5 show the results of E7 and E10 on both novices as representative of the data gathered on all six experts. Although there was not complete overlap in the comments by the experts on both novices, commonality existed. Expert 7 had two comments for each on focusing on the problem (Table 2: #1 and 2 and Table 3: #1 and 2) and one for each on the need to talk to others (Table 2, #6 and Table 3: #5), of needing direct information sources (Table 2: #9 and Table 3: #6), the necessity of multiple sources (Table 2: #5 and Table 3: #7) and the need to know your audience (Table 2: #11 and Table 3: #11). Expert 10 had two comments for trainee 1 (Table 4: #3 and #6) and one for trainee 2 (Table 4: #3) on using metadata in titles; one comment each for trainee 1 and for trainee 2 on keeping task focus (Table 4, #1 and Table 5: #4), looking for independent reviews (Table 4: #11 and Table 5: #9), using several sources (Table 4: #12 and Table 5: #8) and on including implications in the assessment (Table 4: #19 and Table 5: #11).

Table 2. Critiquing Comments of E7 on Novice 1

Analyst	Information Selection	Corroborate/Resolution	Story Construction
E7 on #1	<ol style="list-style-type: none"> 1. Understand what the question really is 2. Have a goal in mind 3. Need to check intelligence sources not just open source 	<ol style="list-style-type: none"> 4. Be aware of directed sources, where they only put in what they want you to believe 5. Need multiple sources to confirm data 6. Talk to other people to get their take 7. Reports six months or so after an event (depending on the event) probably have more accurate information than those immediately around event 8. Take open source with a grain of salt - might be on soap box, misled themselves, intentionally misleading audience 9. Human sources have to have direct knowledge for credibility 	<ol style="list-style-type: none"> 10. Get familiar before formulating 11. Need to know the audience is very true 12. Senior people want pictures, graphs 13. Include basics if the audience might not be fully informed 14. Admit what you don't know 15. You have to be right with what you say 16. Never contradict yourself

Table 3. Critiquing Comments of E7 on Novice 2

Analyst	Information Selection	Corroborate/Resolution	Story Construction
E7 on #2	<ol style="list-style-type: none"> 1. Understand the problem 2. Keep full task in mind as you read 3. Look for patterns in the title list 4. Don't build results on one search. 	<ol style="list-style-type: none"> 5. Talk to other analysts to see to discuss the problem 6. What source information comes from is very important, loses validity if 2nd or 3rd hand information 7. It's necessary to corroborate information, might not use if only in one source. 	<ol style="list-style-type: none"> 8. Carefully construct story 9. If you mention something, be prepared to fully explain 10. Senior people want few words and lots of pictures and graphs in a summary version 11. Know the audience 12. Management wants to know the implications

Table 4. Critiquing Comments of E10 on Novice 1

Analyst	Information Selection	Corroborate/Resolution	Story Construction
E10 on #1	<ol style="list-style-type: none"> 1. Can get wrong focus if don't correctly define problem 2. Go back to the person asking question to resolve ambiguity 3. If using on-screen views of title list, make sure window is large enough to notice patterns of large part of data set 4. Read titles to get a feel for the data set 5. Value of data dependent upon source 6. Use meta data in title list, that is, for example, the fact one name shows up a multitude of times 	<ol style="list-style-type: none"> 7. Ask other analysts for opinions 8. Will have to evaluate data, not just read it 9. Language can be considered a high level determinant during data evaluation 10. Watch for biased reporting 11. Be aware of information that is actually only repeating another source 12. Don't base assessment on only a few documents 13. Be aware complete details might intentionally not be revealed 14. Be aware information might not be directly stated 	<ol style="list-style-type: none"> 15. Look for patterns 16. Sometimes setting down a timeline helps 17. Use external memory, if necessary, to jog memory 18. Diagram the info if possible 19. Include implications in assessment 20. Have list of sources to support assessment 21. Accept that time and resource constraints affect the possible result

Table 5. Critiquing Comments of E10 on Novice 2

Analyst	Information Selection	Corroborate/Resolution	Story Construction
10 on #2	<ol style="list-style-type: none"> 1. Do more than one search 2. Look for implications 3. Use information in titles but don't fully judge documents with that 4. Keep focused on task 5. Look for information over time 	<ol style="list-style-type: none"> 6. Compare information over time to look for changes. 7. Always have to evaluate sources. 8. Use several sources 9. Look for independent reviews 	<ol style="list-style-type: none"> 10. Be prepared for in-depth questions 11. Include implications.

DISCUSSION

Reviewing the protocols indicated that critiquing did address the challenges of proper grounding in context, limited accessibility to experts and tasks, being labor intensive and time consuming, and difficulties with repeatability. Grounding in context is important to encourage comments that are true and interesting for the domain being investigated. The task and the domain used for this initial investigation into the critiquing methodology allowed uncomplicated means for grounding in context. Task accessibility issues are addressed by the critiquing method used in this study as once the novice's process was captured, the re-created novice performance could and was used multiple times. Physical access to experts was overcome because the critiquing session was held outside of normal work areas while still accessing the tool set used within their normal environment. Access challenges because of the existence of only a few experts could be overcome with this method if several novices were used to create different presentations and the experts critiqued the different novices. Another access issue that arises in some domains is that experts are reluctant to participate due to repercussions, such as erroneously performing an act while being observed that could have dire consequences. For critiquing, as the expert is in a role other than performing work, he may see the elicitation as less of a threat and therefore be willing to participate. None of the experts who participated in this study as critiquers were reluctant to participate.

All CTA methods are time consuming, especially the knowledge elicitation phase. One original belief for this study was that by presenting the packaged representation of the novice's process, the experts would use less time to comment than would have been used by having the experienced analysts directly observe the novice. Surprisingly, the novice took about 1 hour and 20 minutes and the experienced analysts ranged from 1 hour and 10 minutes to 1 hour and 21 minutes. Although the times were about the same, the experienced analysts did not have to sit through the less informative times, such as waiting for query results, so the time they spent in their sessions was all productive time focused on the task. Therefore, based on this research, the critiquing method has some increased efficiencies.

Repeatability of the method was demonstrated in this study as the same presentation was shown to all six different experts and all experts made insightful comments. A total of 133 relevant comments were made with the range being 17 to 48. The comments were evaluated to determine if there was an obvious breakdown by year's experience, prior knowledge or work experience and no trends were found. This result is not surprising considering the small sample size. Repeatability was also indicated as E7 and E10 made similar comments when they critiqued the first novice as when they critiqued the second novice as discussed above. The results of all eight critiques had similarities.

While there is not a direct relationship between the tasks investigated for intelligence analysis and tasks performed by commanders, similarity exists. Strategy development entails the steps of mission analysis, course of action development, course of action analysis, course of action comparison and course of action selection. At a high level, these can be compared to the steps previously discussed of information selection, corroboration/resolution and story construction as the commander is gathering and sorting through available information on blue, red and gray forces, comparing inputs to ensure his understanding is correct and then putting together the appropriate course of action to address the objectives based on his understanding of the complete situation. The next step in this research, then, would be to compose an hour-long simulation of a commander in a risky, fast-paced situation as the scenario to be critiqued. An experienced commander, possibly one of the Air Force's mentors, would then be shown the scenario and his comments captured. As with this study, the commander would be encouraged to comment on what was done well, what was done poorly and what he would have done given the same situation.

REFERENCES

- Fischer, G., Lemke, A. C., and Mastaglio, T. (1991). Critics: an emerging approach to knowledge-based human-computer interaction. *International Journal of Man-Machine Studies*, Vol 35: 695-721.
- Hoffman, R. R., Shadbolt, N. R., Burton, A. M., and Klein, G. (1995). Eliciting knowledge from experts: a methodological analysis. *Organizational Behavior and Human Decision Processes* 62 (2), 129-158
- Hoffman, R. R., Crandall, B., and Shadbolt, N. (1998). Use of the critical decision method to elicit expert knowledge: a case study in the methodology of cognitive task analysis. *Human Factors* 40(2), 254-276.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Klein, G., Calderwood, R. and Clifton-Cirocco, A. (1986). Rapid decision making on the fire ground. *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting*. Dayton, OH. 576-580.
- Lesgold, A. M., Robinson, H., Feltovich, P., Glaser, R., Klopfer, D., and Wang, Y. (1988). Expertise in a complex skill: diagnosing x-ray pictures. In M. T. H. Chi, R. Glaser, and M. Farr (Eds), *The Nature of Expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McNeese, M. D., Zaff, B. S., Citera, M., Brown, C. E. and Whitaker, R. (1995). AKADAM: Eliciting user knowledge to support participatory ergonomics. *International Journal of Industrial Ergonomics*, Vol 15, 345-363.

Potter, S. S., Roth, E. M., Woods, D. D. and Elm, W. (2000). Bootstrapping Multiple Converging Cognitive Task Analysis Techniques for System Design. In Schraagen, J.M.C., Chipman, S.F., & Shalin, V.L. (Eds.), *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sarter, N. B. and Woods, D. D. (2000). Teamplay with a powerful and independent agent: a full-mission stimulation study. *Human Factors* 42(3), 390-402

Woods, D. D. (1993). Process tracing methods for the study of cognition outside of the experimental psychology laboratory. In G. Klein, J. Orasanu, and R. Calderwood (Eds.), *Decision Making in Action: Models and Methods*. (pp. 228-251) Norwood, NJ: Ablex Publishing Corporation,

AUTHOR BIOGRAPHY

JANET E. MILLER is a Research Scientist with an emphasis in decision sciences at the Air Force Research Laboratory, Human Effectiveness Directorate, Wright Patterson AFB, Dayton, OH. Her experience includes gathering user requirements and implementing automated systems at the Pacific Air Force and the National Air and Space Intelligence Center. Her degrees include a BS in Electrical Engineering from Louisiana Technical University (1984), an MS in Management of Information Systems from the University of Arizona (1988) and a Ph. D. in Cognitive Systems Engineering (2002) from the Ohio State University. Additionally, she is a LtCol in the US Air Force Reserves (Active).

DISCLAIMER

The views expressed in this paper are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense or the U. S. Government.

Critiquing: A Methodology to Extract C2 Expertise

18 June 2003



Dr. Janet E. Miller
Cognitive Systems Engineer
Human Effectiveness Directorate
Air Force Research Laboratory



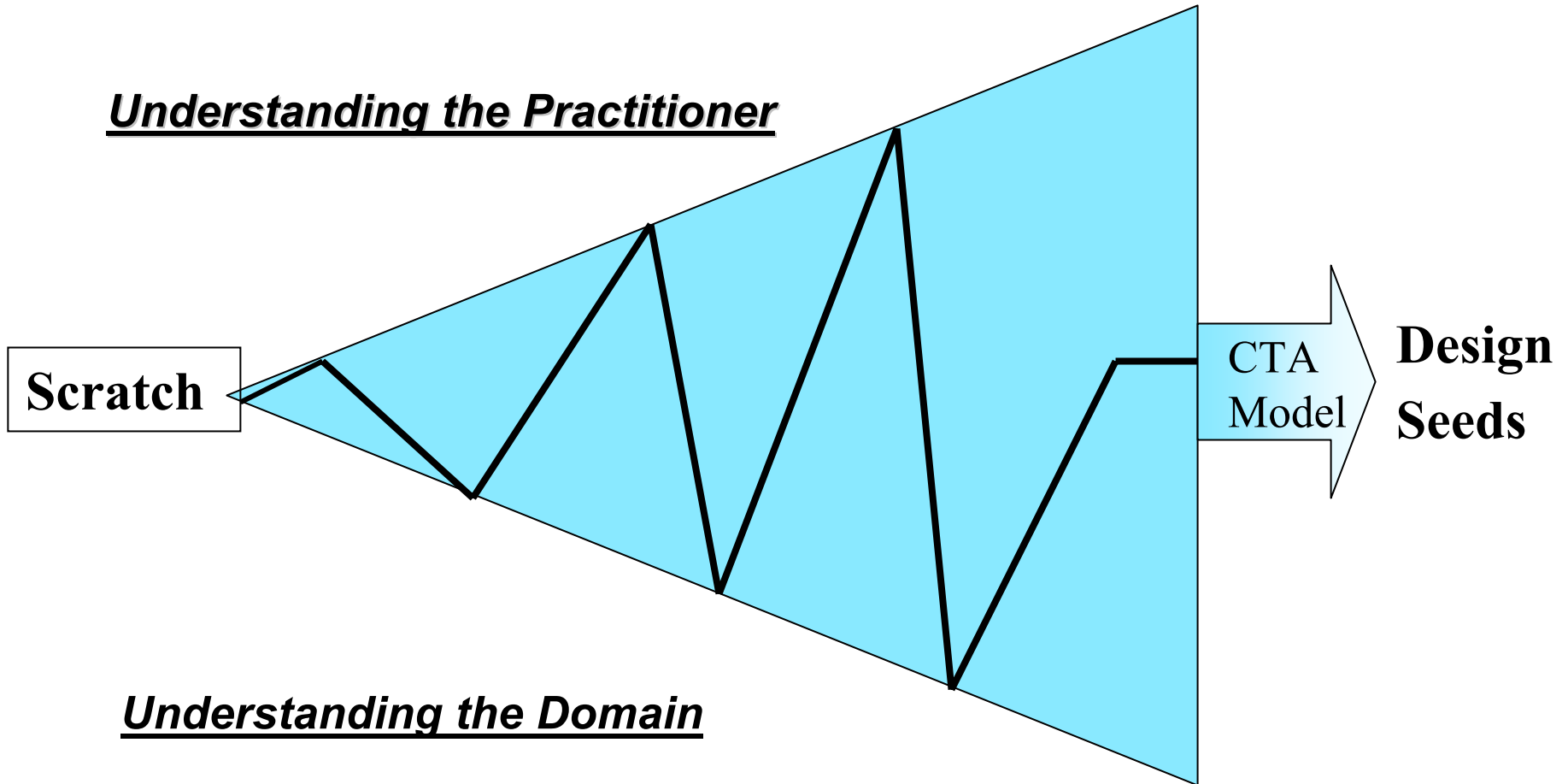
Problem



- **In-depth understanding of C2 is required for developing accurate, relevant training and support**
- **Extracting tacit knowledge is difficult**
 - **Few experts**
 - **Unpredictability**
 - **Often joint, international positions**
- **Existing knowledge elicitation methods have challenges**
 - **Grounding in context**
 - **Accessibility to experts and relevant tasks**
 - **Laborious and time consuming**
 - **Repeatability**



Gaining Understanding



Exploring the current world → Exploring the future world



Process of Expert Performing Critiquing Task



Novice

Performing

**Actual/Simulated
Task**

Process of Novice Performing Domain Task

**CTA
Investigator**

Watches

**Expert
as Evaluator**

Commenting on

Novice Performance



Beginning Note Sheet



****EXERCISE EXERCISE EXERCISE EXERCISE****

In 1996, the European Space Agency lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.



Complete Novice Note Sheets



****EXERCISE EXERCISE EXERCISE EXERCISE****

Ariane 5 --> replacing
Launcher A.4.

flight
--> Ariane 501

Russian Rocket dead -
Same Week

In 1996, the [European Space Agency] lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.

5 Jun 1996

over-
confidence

SW failure
Blow up

exclusively

did not carry
out tests on
IRS
--> could have
found problem

errors in
inertial reference
system (IRS,
backup failed, too)

Loss of guidance
30 secs > L.O.
Got bad info, thought
it was
(2.5 miles. Self destruct)

Ariane 5
--> SW failure?
Blew up

No
Expected
Problem

Sunstorm
[-prevent EM
storms on
sun

Inquiry
Board --> ESA
CNES (French)

implemented

- IRS should be turned
·off immediately after lift off
- [test facility]☆ --> more thorough testing program
for entire launcher
 - ↳ w/as much real equipment
as technically feasible

- More transparent work/w Ariane 5 partner
- 384\$ recovery from failure

Next
flight
mid April
1997

get well program

unplanned

Lost
year

what they did:

- rewritten sw
- > in light of lessons

Made better testbed
analyzed all SW

--> used A-4
SW

fill in
w/old
tech

when

-5 Jun 1996, Ariane 5 program, replacing Ariane 4 program

why

SW failure-IRS system malfunction(wrong system: for A-4 program
didn't expect problems

--> did not test

--> could have prevented

Impacts--> \$384 blunder, lost year of program
unplanned flight

fill in Ariane 4
rocket

get well program
re-analyze all sw (IRS), made more realistic sw



Progressive Novice Note Sheets



****EXERCISE EXERCISE EXERCISE EXERCISE****

Ariane 5 --> replacing
Launcher A.4.

flight
--> Ariane 501

Russian Rocket dead -
Same Week

In 1996, the [European Space Agency] lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.

5 Jun
1996

Ariane 5
--> SW failure

****EXERCISE EXERCISE EXERCISE EXERCISE****

Ariane 5 --> replacing
Launcher A.4.

flight
--> Ariane 501

Russian Rocket dead -
Same Week

In 1996, the [European Space Agency] lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.

5 Jun
1996

Ariane 5
--> SW failure ?
Blew up



Progressive Novice Note Sheets



****EXERCISE EXERCISE EXERCISE EXERCISE****

Ariane 5 --> replacing
Launcher A.4.

Russian Rocket dead -
Same Week

flight
--> Ariane 501

In 1996, the [European Space Agency] lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.

5 Jun 1996

SW failure
Blow up

Ariane 5
--> SW failure?
Blew up

****EXERCISE EXERCISE EXERCISE EXERCISE****

Ariane 5 --> replacing
Launcher A.4.

Russian Rocket dead -
Same Week

flight
--> Ariane 501

In 1996, the [European Space Agency] lost a satellite during the first qualification launch of a new rocket design. Give a short briefing about the basic facts of the incident: when it was, why it occurred, and what the immediate impacts were.

5 Jun 1996

SW failure
Blow up

Ariane 5
--> SW failure?
Blew up

Sunstorm
[-prevent EM
storms on
sun



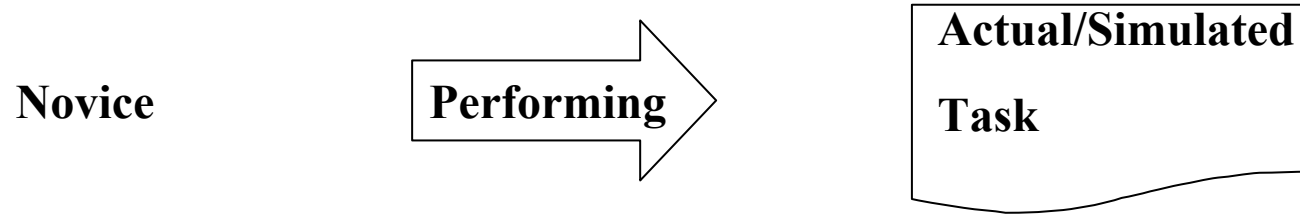
Experts' Comments about Study



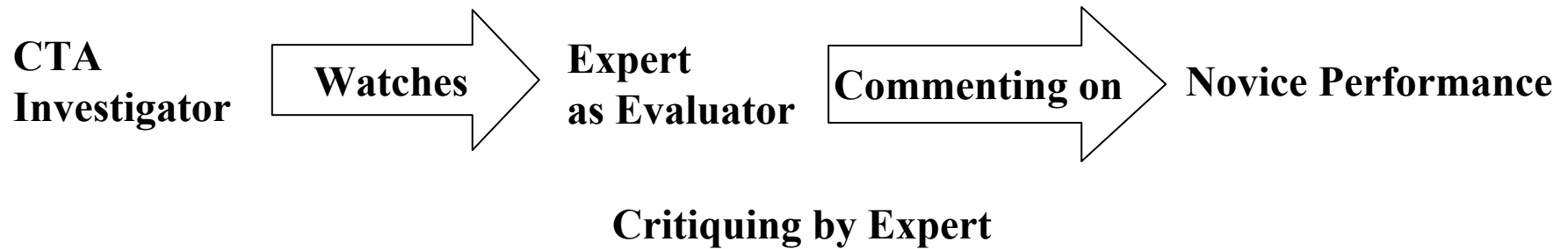
- E5 – No prior knowledge:** I felt uncomfortable
- E6 – No prior knowledge:** I was a little uncomfortable.
- E8 – Partial prior knowledge:** I'm glad I was prepared. I was able to draw my own conclusions.
- E9 – Partial prior knowledge:** I would not have been comfortable without at least reading the articles, seeing what data set he's working with. If you don't know what data set they have, you can't critique, it's not valid.
- E7 – Full prior knowledge:** I was comfortable in the way data was presented. You might tell them (novices) but they have to do it themselves. This is good for training as they do the work and not just memorize.
- E10– Full prior knowledge:** I had fun. I found it intriguing.



Process of Expert Performing Critiquing Task



Process of Novice Performing Domain Task





Discussion



- **Remember: Any method shapes the conditions of observation.**
- **Relationship to C2**
 - **Can be used in conjunction with Modeling & Simulation**
 - **Provides strong cue to focus**
 - **Participatory role**



Conclusion



- **Existing knowledge elicitation methods have challenges**
 - **Grounding in context**
 - **Accessibility to experts and relevant tasks**
 - **Laborious and time consuming**
 - **Repeatability**
- **Critiquing used to get an in-depth understanding of the C2 domain required for developing accurate, relevant training and support**

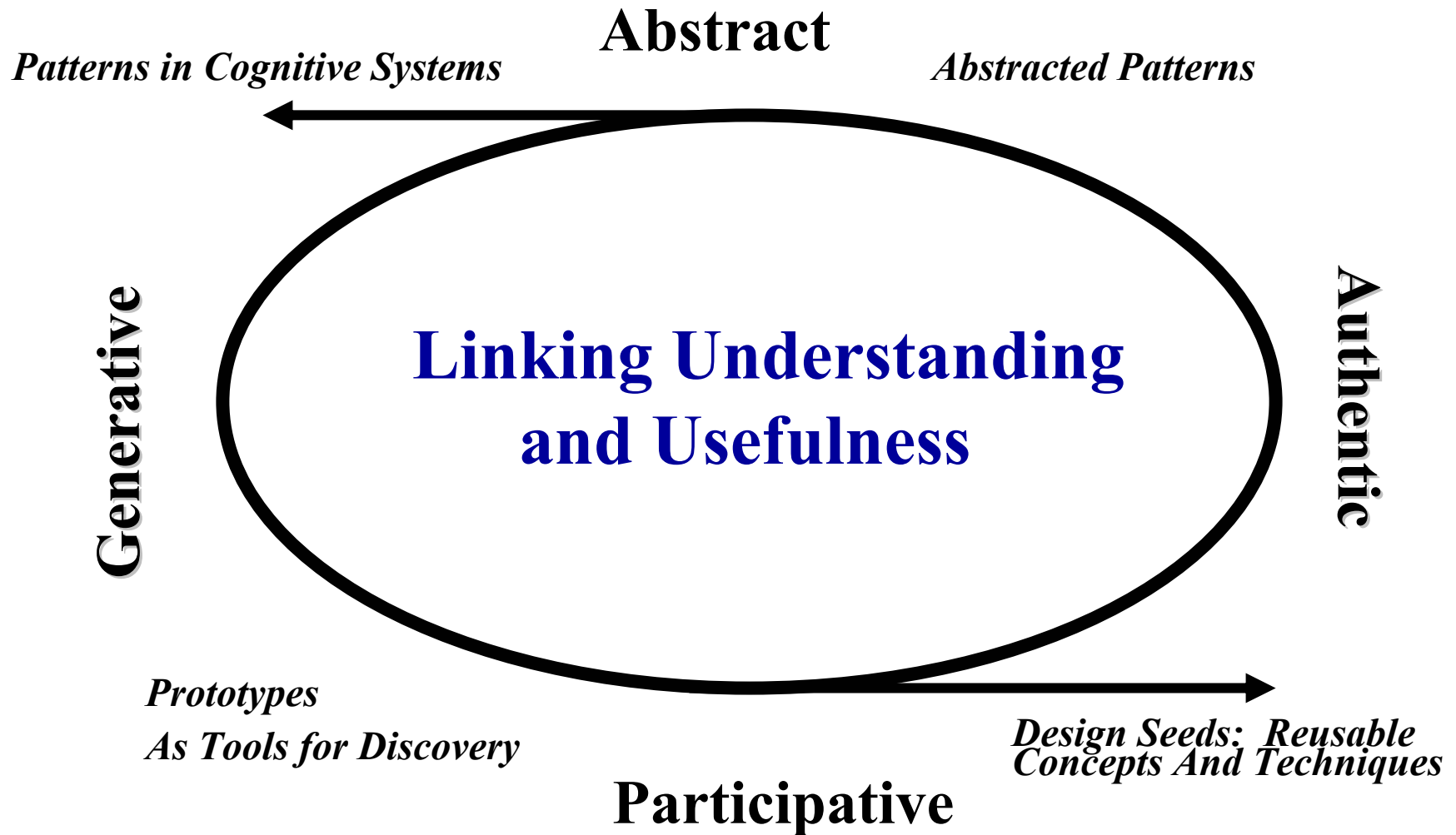


Backups





Designing Tools for People





Method Breakdowns



Families:

- **Structured Interview/Elicitation**
- **Performance/Observation**
- **Mapping Semantic Space**
- **Functional Domain Modeling**
- **Functional Task Modeling**
- **Computational Modeling Methods**
- **Participatory Design Methods**

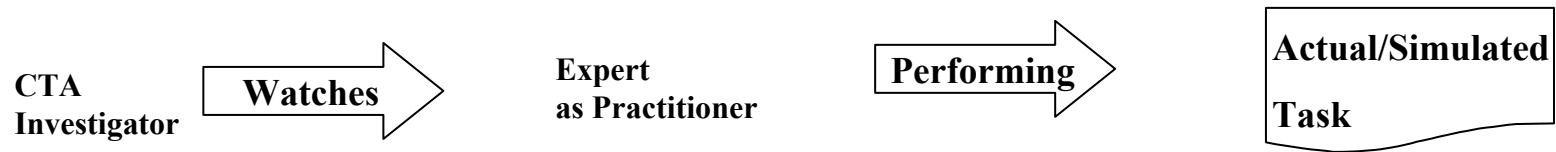


Comparison of Elicitation Methods

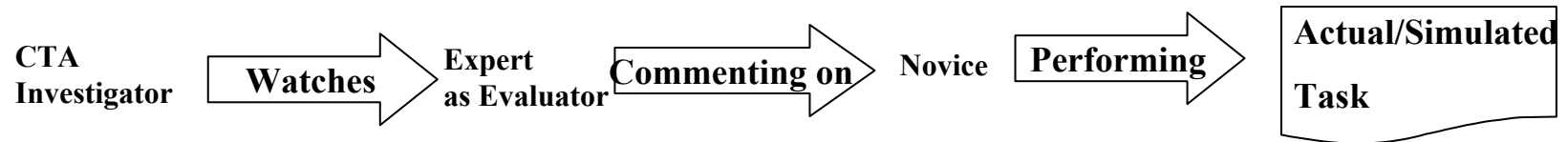
Elicitation by Interview



Elicitation by Observation



Elicitation by Critiquing





Analyst E7 on Novice 1



•Information Selection:

- Understand what the question really is
- Have a goal in mind
- Need to check intell sources not just open course

•Corroborate/Resolution:

- Be aware of directed sources
- Need multiple sources to confirm data
- Talk to other people
- Reports 6 months or so after an event probably more accurate
- Take open source with a grain of salt
- Human sources have to have direct knowledge

•Story Construct:

- Get familiar before formulating
- Need to know the audience
- Senior people want pictures, graphs
- Include basics if audience might not be fully informed
- Admit what you don't know
- You have to be right with what you say
- Never contradict yourself



Analyst E7 on Novice 2



•Information Selection:

- Understand the problem
- Keep full task in mind as you read
- Look for patterns in the title list
- Don't build results on one search.

•Corroborate/Resolution:

- Talk to other analysts to see to discuss the problem
- What source information comes from is very important, loses validity if 2nd or 3rd hand information
- It's necessary to corroborate information, might not use if only in one source

•Story Construct:

- Talk to other analysts to see to discuss the problem
- What source information comes from is very important, loses validity if 2nd or 3rd hand information
- It's necessary to corroborate information, might not use if only in one source



Analyst E10 on Novice 1

Information Selection:

- Can get wrong focus if don't correctly define
- Go back to the person asking question to resolve ambiguity
- If using on-screen views of title list, make sure window is large enough to notice patterns of large part of data set
- Read titles to get a feel for the data set
- Value of data dependent upon source
- Use meta data in title list, that is, for example, the fact one name shows up a multitude of times

Corroborate/Resolution:

- Ask other analysts for opinions
- Will have to evaluate data, not just read it
- Language can be considered a high level determinant during data evaluation
- Watch for biased reporting
- Be aware of information that is actually only repeating another source
- Don't base assessment on only a few documents
- Be aware complete details might intentionally not be revealed
- Be aware information might not be directly stated

Story Construct:

- Look for patterns
- Sometimes setting down a timeline helps
- Use external memory, if necessary, to jog memory
- Diagram the info if possible
- Include implications in assessment
- Have list of sources to support assessment
- Accept that time and resource constraints affect the possible result



Analyst E10 on Novice 2



• Information Selection:

- Do more than one search
- Look for implications
- Use information in titles but don't fully judge documents with that
- Keep focused on task
- Look for information over time

• Corroborate/Resolution:

- Compare information over time to look for changes.
- Always have to evaluate sources.
- Use several sources
- Look for independent reviews

• Story Construct:

- Be prepared for in-depth questions
- Include implications.